

국소 극대-극소점간의 간격정보를 이용한 시간영역에서의 음성인식을 위한 파라미터 추출 방법

(A Time-Domain Parameter Extraction Method for Speech Recognition using the Local Peak-to-Peak Interval Information)

林材烈*, 金亨一**, 安秀桔*

(Jae Yeol Rheem, Hyong Ill Kim, Beum Hun Kim and Sou Guil Ann)

要約

본 논문에서는 음성인식을 위한 시간영역에서의 새로운 파라미터 추출방법을 제안한다. 제안된 방법에서는 음성 파형에서 국소 극대-극소점간의 간격정보가 주파수 성분과 밀접한 관계가 있음을 이용하여, 간격정보의 분류를 통한 시간영역에서의 필터뱅크(filter bank)기법을 이용하여 파라미터를 추출한다.

제안된 파라미터 추출방법의 인식시스템에의 이용 가능성을 시험하고자 VQ(Vector Quantization) 및 HMM(Hidden Markov Model)을 이용한 고립단어 인식시스템을 구성하여, 10명의 화자가 22개의 단어를 세번씩 발음한 테스트 데이터에 대하여 인식실험을 수행하였다. 그 결과 92.9%의 인식율을 얻음으로써 제안된 파라미터의 유용성을 확인할 수 있었다. 제안된 방법은 시간영역에서 파라미터를 직접 추출함으로써, A/D 변환기만 있으면 별도의 DSP(Digital Signal Processing)보드 없이 개인용 컴퓨터에서 실시간 추출이 가능하다.

Abstract

In this paper, a new time-domain parameter extraction method for speech recognition is proposed. The suggested method is based on the fact that the local peak-to-peak interval, i.e., the interval between maxima and minima of speech waveform is closely related to the frequency component of the speech signal. The parameterization is achieved by a sort of filter bank technique in the time domain.

To test the proposed parameter extraction method, an isolated word recognizer based on Vector Quantization and Hidden Markov Model was constructed. As a test material, 22 words spoken by ten males were used and the recognition rate of 92.9% was obtained. This result leads to the conclusion that the new parameter extraction method can be used for speech recognition system. Since the proposed method is processed in the time domain, the real-time parameter extraction can be implemented in the class of personal computer equipped only with an A/D converter without any DSP board.

1. 서론

*正會員, **學生會員, 서울대학교 電子工學科
(Dept. of Elec. Eng., Seoul Nat'l Univ.)

※본 연구는 93년도 삼호전자의 연구비 지원으로
수행되었습니다.

接受日字 : 1993年 4月 26日

음성을 매개체로 하는 인간과 기계간의 통신을 위하여 지난 20여년 동안 음성인식을 위한 여러 가지 방법이 연구되어져 왔다. 그 결과 오늘날에는 HMM (Hidden Markov Model)과 신경회로망(Neural

Network) 방법 등을 이용한 인식시스템이 주류를 이루고 있으며, 이러한 인식시스템에 사용되는 파라미터로서는 성도모델에 근거한 LPC계수, cepstrum 계수, 그리고 필터뱅크의 출력 등을 들 수 있다.^[1] 그런데 이러한 파라미터들은 모두 복잡한 처리단계를 거쳐서 주로 주파수 영역의 파라미터로서 추출되기 때문에, 추출된 파라미터를 이용한 인식과정의 처리는 별도로 치더라도, 음성 파형으로부터 파라미터를 추출하는 과정을 실시간으로 처리하기 위해서는 빠른 처리속도의 하드웨어가 요구된다.

시간영역에서 음성 파형으로부터 직접 인식용 파라미터를 추출할 수 있다면, 주파수 영역에서의 처리에 따른 복잡한 계산이 필요 없기 때문에 일반적으로 빠른 추출이 가능해진다. 지금까지 시간영역에서의 파라미터화에 의한 음성인식 연구결과를 살펴보면, 음성신호처리의 기본적인 파라미터인 영교차율 등을 이용한 경우^[2,4]와 파형의 특징을 추출하여 통계적으로 처리한 경우^[5] 등으로 극히 제한적이었으며, 모두 본격적인 인식시스템으로 개발되지는 못하였다.

본 논문에서는 시간영역에서 새로운 인식용 파라미터의 추출방법을 제안한다. 제안된 방법은 음성 파형에서 국소 극대-극소점간의 간격정보가 주파수 성분과 밀접한 관계가 있다는 사실에 근거하여, 간격정보에 대한 일종의 필터뱅크 기법을 이용하여 파라미터를 추출하고자 한다. 제안된 파라미터 추출방법의 유용성을 확인하고자 소규모 어휘의 고립단어 인식 시스템을 구현한 결과, 92.9%의 인식율을 얻음으로써 인식시스템에 이용가능함을 확인할 수 있었다. 제안된 파라미터 추출방법은 A/D 변환기가 있는 경우에 별도의 추가 장비 없이 개인용 컴퓨터에서도 실시간 추출이 가능하여, 개인용 컴퓨터에서의 인식시스템의 구현에 활용도가 높음을 알 수 있다.

II. 제안된 파라미터 추출방법

1. 국소 극대-극소점간의 간격과 주파수 성분

음성 파형을 극단적으로 왜곡시켜서 만든 파형을 이용하여 인지실험을 한 것을 살펴보면, 음성의 인지를 유지하는 중요한 요소가 국소 극대-극소점간의 간격(local peak-to-peak interval, i.e., interval between maxima and minima)임을 알 수 있다.

^[6] ^[7] Licklider와 Pollack의 인지실험을 살펴보면, 원래의 음성 파형을 미분하고 극단적으로 클리핑시켜서 만든 음성 파형의 인지도 값(intelligibility score)이 97%로써 원래의 음성 파형이 가지고 있는 인지도 99%에 비하여 인지도의 저하가 거의 일어나지 않음

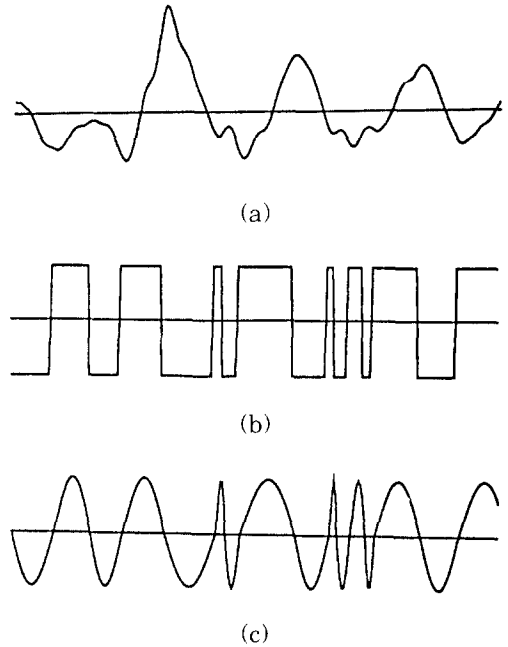


그림 1. 왜곡된 파형

(a) 원래의 파형 (b) 미분 및 클리핑된 파형 (c) 사인함수 형태로 재구성한 파형

Fig. 1. Amplitude distorted waveform.

(a) original waveform, (b) differentiated and clipped waveform, (c) reconstructed waveform with sine shape.

을 알 수 있다.^[6] 원래의 파형을 미분하여 극단적으로 클리핑한 경우에는 그림 1의 (b)와 같이 일정한 크기를 가진 펄스열이 되며, 펄스열의 극성이나 모양이 음성의 인지도에 영향을 크게 미치지 않는다고 알려져 있다.^[7] 반면에 영교차 간격(zero-crossing interval)이 유지되게 클리핑한 파형의 인지도는 70%로써 그 인지도가 크게 떨어진다. 따라서 과거에 영교차 간격을 이용한 파라미터화 방법들^{[2], [3]}은 인지도가 보장되지 않는 파라미터임을 알 수 있다.

음성 파형을 미분하여 극단적으로 클리핑시켰을 때 남아있는 유일한 정보는 극대-극소점간의 간격이므로, 인간의 음성 인지작용이 청각계의 주파수 처리에 의존한다는 사실에 비추어 보면, 극대-극소점간의 간격정보가 주파수 성분과 밀접한 관계가 있음을 알 수 있다. 그리고 음성 파형의 왜곡실험을 확장하여, 간격정보를 유지한 채 각각의 크기정보를 원래의 파형 크기에 맞추어 조절하게 되면, 음성이 지닌 개인성 정보가 살아남을 관찰할 수 있어 음성신호를 간격 및

크기정보로 나타낼 경우, 간격정보에 주파수 정보가 크기정보에 개인성 정보가 들어있음을 알 수 있다. 그림 1의 (c)를 살펴보면, 간격정보가 가지는 주파수 성분은 반주기가 간격정보에 해당하는 주파수라고 가정할 수 있다. 균일 표본화된 신호에서는 간격정보가 샘플수로 나타나므로, 표본화 주파수 F_s 로 균일 표본화된 신호에서 간격정보의 길이가 i 샘플일 경우에 그 간격정보에 해당하는 주파수 성분 f_i 는 표본화 간격을 $\Delta T(=1/F_s)$ 라 할 때, 다음과 같이 표현된다.

$$f_i = \frac{1}{2 \cdot i \cdot \Delta T} = \frac{F_s}{2 \cdot i} \quad (1)$$

원래의 파형을 국소 극대-극소점간의 간격을 고려하여 그림 1의 (c)와 같이 사인함수 형태의 펄스열로 재구성한 후 신호의 스펙트럼을 비교하여 보면 그림 2에서와 같이 포오만트 정보가 유지됨을 관찰할 수 있으며, 이로써 위의 가정이 타당함을 알 수 있다. 그림 3에는 표본화 주파수가 42kHz일 때, 식 (1)에 의한 간격정보의 길이와 주파수성분 관계를 보여준다.

2. 파라미터 추출방법

간격정보와 그 주파수성분 관계를 이용하면, 시간 영역에서 국소 극대-극소점간의 간격정보를 추출하여 그 분포를 분석함으로써 시간영역에서의 주파수 영역 분석이 가능하게 된다. 주파수 성분 분석은 대상신호가 가지는 주파수 영역을 몇 개의 대역으로 나누고

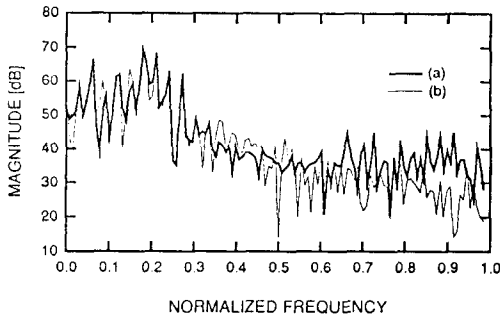


그림 2. 모음 /아/의 스펙트럼 비교
(a) 원래 음의 FFT 스펙트럼
(b) 사인함수 형태로 재구성한 파형의 FFT 스펙트럼

Fig. 2. Spectrum comparison of Vowel /a/.
(a) FFT spectrum of original /a/.
(b) FFT spectrum of reconstructed signal of (a) with sine shape.

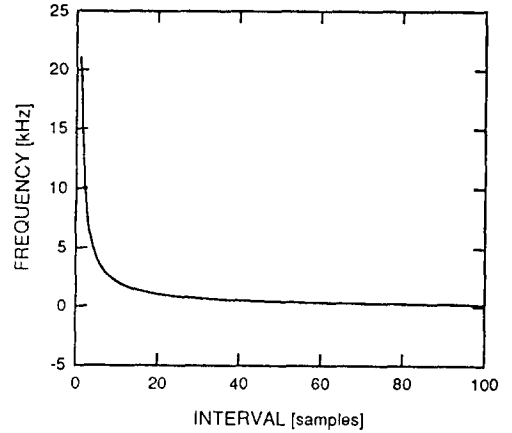


그림 3. 42kHz 표본화시의 간격과 주파수 성분 관계
Fig. 3. Relation between interval and its frequency component at 42kHz sampling.

일정시간 동안의 각 대역필터의 출력력을 파라미터로 이용하는 필터뱅크 분석기법을 적용하여, 시간영역에서 각각의 대역에 해당하는 간격정보의 갯수를 정규화하여 출력함으로써 얻어질 수 있다.

음성대역을 L 개의 대역필터로 나누고, 각각의 대역필터의 주파수 영역집합을 $F_k, k=1, 2, \dots, L$ 로 나타내자. 그리고 j 번째 프레임에서 N_j 개의 간격정보가 존재하고 이중 k 대역에 속하는 간격정보의 갯수를 $n_{kj}, k=1, 2, L, j=1, 2, \dots$ 라 하면, 각 대역의 출력 P_k 는 다음과 같이 정의된다.

$$P_j = \frac{1}{N_j} n_{kj}, \quad K=1, 2, \dots, L \text{ and } j=1, 2, \dots \quad (2)$$

$$\text{where } \begin{cases} N_j = \sum_{k=1}^L n_{kj} \\ n_{kj} = \text{no. of interval } i, \text{ for } f = \frac{F_s}{2 \cdot i} \in F_k, k=1, 2, \dots, L. \end{cases}$$

이때 j 번째 프레임의 시간축상 길이 T_j 는

$$T_j = N_j \cdot \Delta T = \sum_{k=1}^L n_{kj} \cdot \Delta T \quad (3)$$

가 된다. 이러한 간격정보의 분포 $\{P_{kj}\}, k=1, 2, \dots, L$ 는 시간에 따른 간격정보의 분포를 나타내며 이것을 음성인식용 파라미터로 사용한다.

식 (1)을 살펴보면, 간격과 주파수의 관계가 기본적으로 표본화 주파수에 의하여 제약을 받으므로, 시간영역에서 간격정보를 이용하여 주파수성분 분석을

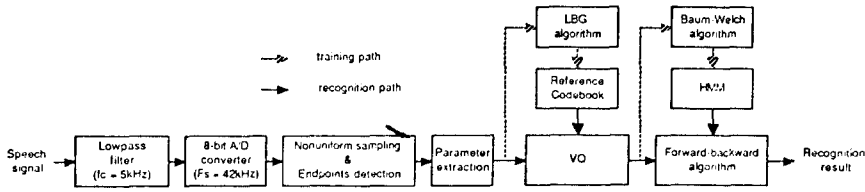


그림 4. 고립단어 인식시스템의 구성도

Fig. 4. Block diagram of isolated word recognition system.

할 경우에는 표본화 정리(sampling theorem)^[9] 보다 높은 주파수로 오버 샘플링하는 것이 필요하다. 표본화 주파수가 낮은 경우에는, 식 (1)의 가정에 의하여 관찰할 수 있는 주파수 성분 자체가 제한되어, 전체 음성대역에서 관찰할 수 있는 주파수 성분의 개수가 줄어들게 된다. 반면에 높은 주파수로 오버 샘플링하는 경우에는 간격정보의 분해능이 높아져서 세밀한 주파수 성분 분석이 가능하며 전체 음성대역에서 관찰할 수 있는 주파수 성분의 개수가 많아지게 되어 개별발음에 대한 변별력이 커지게 된다. 그리고 이러한 파라미터 값들은 원래의 신호가 가지고 있는 다양한 주파수 정보중 언어적 정보의 전달에 필요한 충분한 주파수 분포 정보를 지니고 있어서 개별발음마다 고유의 변별성과 일관성을 보여주기 때문에 음성인식에 사용할 수 있다.

III. 실험 및 결과

제안된 파라미터화 방법으로 추출된 파라미터의 음성인식에 대한 유용성을 테스트하기 위하여 VQ 및 HMM을 이용한 소규모 어휘의 고립단어 인식시스템을 그림 4와 같이 구성하였다.

인식 대상어휘는 숫자음과 컴퓨터의 조작에 필요한 것이라고 추정되는 명령어를 포함한 22개의 단어이며, 20대 남성화자 10명이 잡음환경하의 컴퓨터실에서 세번씩 발성한 것을 직접 A/D 변환하여 학습에 사용하였다. 마이크로 입력된 음성신호는 대역폭 5kHz의 저역필터를 통과하여 A/D 변환되며, A/D 변환시에는 표본화 주파수 F_s 가 42kHz로써, 약 4배의 오버샘플링을 취하여 간격 정보의 분해능을 높였다. 표 1은 인식실험에 사용한 단어목록이다.

표 1. 인식대상어휘

Table 1. Words used for recognition.

숫자	공, 일, 이, 삼, 사, 오, 육, 칠, 팔, 구, 십
명령어	좌로, 우로, 위, 아래, 선택, 탈출, 처음, 끝, 다음, 예, 아니오

간격정보의 추출에는 최근 제안된 비균일 표본화 알고리즘을 사용하였다.^[8] 비균일 표본화 알고리즘은 국소 극대 및 극소점에서 음성신호를 비균일 표본화하여 표본점에서의 크기와 간격열을 이용하여 음성신호를 나타내는 방법으로, 본 실험에서는 그 정보중 간격열만을 이용했다. 그리고 음성구간의 검출(endpoints detection)은 비균일 표본화 알고리즘 상에서 묵음구간의 경우 일정한 간격정보를 가지게 처리하여 묵음구간을 배제할 수 있도록 했다. 사용된 묵음구간의 처리방법은 다음과 같다.

묵음구간 처리방법

연속한 윈도우를 잡아 윈도우 내의 균일 표본화점의 통계적 특징을 추출하여 묵음으로 판정되면 윈도우 내의 마지막 샘플을 국소 극대 및 극소점과 관계없이 비균일 표본화점으로 취한다. 이때 간격정보는 윈도우의 길이가 된다. 묵음이 아닌 음성구간으로 판정되면 제안된 비균일표본화알고리즘을 적용하여 비균일 표본화를 수행한다. 음성구간에 이어서 묵음구간이 오는 경우에는 음성구간의 마지막 샘플값을 국소 극대 및 극소점에 관계없이 비균일 표본화하여, 다음에 오는 묵음구간의 간격정보를 일정하게 유지시킨다.

묵음구간의 판정에 사용된 파라미터는 윈도우 내의 평균에너지^[10]이며 실험에 사용된 윈도우의 길이는 255이다.

파라미터 추출시에는 음성구간에 해당하는 간격정보열에서 한 프레임을 40샘플씩 ($N_f=40$)잡고 50%씩 중첩하여 추출하였다. 한 프레임을 40 샘플씩 잡은 것은 간격정보의 평균길이를 고려하여 실험적으로 정한 것이다. 파라미터는 음성대역을 10개의 대역으로 나누고 각각의 대역 범위에 포함되는 주파수 성분을 가지는 간격정보의 개수를 프레임내의 전체 간격정보의 수로 정규화하여 추출하였다. 대역의 구분에는 간격정보와 주파수의 관계에서 간격정보를 균등하게 나눌 수 있도록 주파수축에서 정한 것과 Bark scale^[11]을 이용하여 균등대역폭으로 나눈 것 두가지를 사

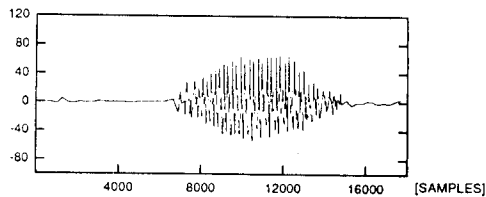
용하였다. 실험에 사용된 필터뱅크의 대역 경계주파수가 표 2에 나타나 있다.

표 2. 필터뱅크의 경계주파수

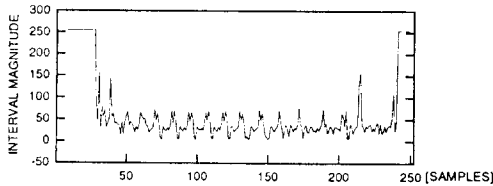
Table 2. Frequency division of filter banks. (단위: Hz)

대역번호	1	2	3	4	5	6	7	8	9	10
균등간격고려	472	532	609	712	857	1077	1448	2211	3231	5000
Bark scale고려	180	374	596	863	1195	1617	2161	2868	3790	5000

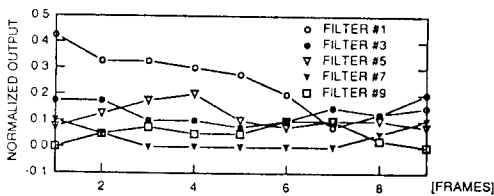
그림 5는 제안된 방법으로 추출된 파라미터 궤적의



(a)



(b)



(c)

그림 5. 제안된 방법으로 추출된 파라미터 궤적 (a) 42kHz로 균일 표본화된 원래의 음성 파형 (b) 비균일 표본화된 간격열 (c) 추출된 파라미터궤적 예

Fig. 5. Parameter contours extracted by the proposed method. (a) original speech waveform uniformly sampled with $F_s = 42\text{kHz}$. (b) interval sequence obtained by nonuniform sampling method (c) parameter contour examples.

한 예를 보여준다. 그림 5의 (a)는 42 kHz로 표본화된 /오/음의 파형이며, (b)는 (a)를 비균일표본화하여 얻은 간격열이다. 간격열에는, 파형에서 나타나는 유성음의 준주기적 특징이 잘 반영되어 있음을 관찰할 수 있다. 그림 5의 (c)는 제안된 방법으로 추출한 파라미터중 일부의 궤적을 보여주며, 대역별 에너지의 분포가 일관성을 갖음을 관찰할 수 있다.

VQ방법으로는 LBG splitting 알고리즘을 사용했고 거리측정(distance measure)에는 cepstrum방법에서 많이 사용하고 있는 가중유클리디안 거리측정(weighted Euclidean distance measure) 방법을 사용했으며 코드북 사이즈(codebook size)는 128이다.^[12] 대역필터 k 번째로 부터 추출된 파라미터 값에 대한 가중치 $w(k)$ 은 (4)식으로

$$w(k) = 1 + \frac{L}{2} \cdot \sin\left(\frac{\pi}{L} \cdot k\right), k = 1, 2, \dots, L \quad (4)$$

표현되며 여기서 L 은 대역필터의 갯수이다.^[13]

HMM방법으로는 상태수가 각각 4와 6인 left-to-right 모델을 이용하였으며, HMM의 학습에는 Baum-Welch 알고리즘을 이용하였고, 인식시에는 forward-backward 알고리즘을 사용하였다.^{[14] [15]}

표 3에는 인식결과가 있다. 인식을 I은 VQ시에 가중치를 고려한 결과이며, 인식을 II는 가중치를 고려하지 않았을 때의 인식결과이다. 인식결과를 보면, 가중치를 고려하고 HMM의 상태수가 6이며 간격정보를 고려한 필터 뱅크를 사용했을 때가 인식을 92.9%로 가장 높음을 알 수 있다.

표 3. 인식 결과

Table 3. Recognition results.

필터뱅크	간격정보 고려		Bark scale고려	
	4	6	4	6
HMM 상태수				
인식율 I(%)	91.7	92.9	90.0	90.9
인식율 II(%)	80.2	82.9	90.0	89.1

실험결과를 살펴보면, VQ시의 가중치의 존재여부와 사용된 필터뱅크에 따라서 인식율의 차이가 현저함에 따라서 대역필터의 설계 및 그에 따른 가중치에 대한 최적화 연구가 선행되어야 함을 알 수 있다. 학습에 사용된 전체 프레임수를 비교하여 보면, 제안된 방법에 의하여 얻어지는 프레임수가 균일 표본화된 신호에서 기존의 파라미터 추출방법으로 얻어지는 프레임수에 비하여 작게 됨을 관찰할 수 있다. 일반적

으로 VQ 및 HMM을 이용한 인식시스템에서는 학습에 사용된 데이터 양에 따라서 인식율의 차이를 보인다. 이는 사실을 고려하면, 발음횟수를 증가시켜서 학습시에 사용하게되는 데이터 양을 증가시킬 경우 인식율이 증가하게 될 것으로 예상된다.

제안된 시간영역 파라미터를 이용하여 얻은 위의 인식율은 기존의 주파수영역 분석기법 파라미터인 cepstrum 계수를 이용하고, VQ시에 가중치를 고려하고, HMM 상태수가 6일 때의 인식율 96.4%에는 다소 떨어진다. 이것은 제안된 시간영역 추출방법이 언어적 정보의 전달에 필요한 최소정보라고 간주할 수 있는 간격정보만을 이용하였고, 간격정보에 해당하는 주파수 성분이 역으로 단순히 간격정보의 형태만으로는 나타나지는 않을 것이라는 점에 비추어서 제안된 시간영역 파라미터가 거친 분석을 제공했기 때문으로 사려된다. 그러나 파라미터 추출에 걸린 시간을 비교하여 보면, 486 DX2-66 개인용 컴퓨터에서 cepstrum 방법이 단어당 평균 약 0.47초, 제안된 방법은 약 0.06초 걸렸다. 따라서 약 8배 빠른 추출이 가능하며, 별도의 DSP 보드의 도움없이도 전체 인식시간이 1초 미만으로 가능하여 개인용 컴퓨터에서도 실시간 인식이 가능함을 알 수 있다.

V. 결론

본 논문에서는 비균일 표본화시에 얻어지는 간격정보가 인지적 측면에서 반주기에 해당하는 주파수 성분과 관계가 있다고 가정하고, 시간영역에서 필터뱅크 기법을 이용한 파라미터 추출방법을 제안하였다. 그리고 소규모 어휘의 고립단어 인식시스템을 구성하여 제안된 파라미터 추출방법의 유용성을 확인하였다.

제안된 추출방법은 언어적 정보의 전달에 필요한 최소정보라고 간주할 수 있는 간격정보만을 이용하였고, 간격정보에 해당하는 주파수 성분이 역으로 단순히 간격정보의 형태만으로는 나타나지는 않을 것이라는 점에 비추면 거친 분석을 제공함으로써, 결과적으로 기존의 주파수 영역의 분석기법에 의한 파라미터를 이용했을 때보다 다소 인식율이 떨어지는 단점을 지니나, 첫째, 시간영역에서 음성의 파형으로부터 직접추출이 가능함으로써 빠른 추출이 가능하며, 소규모 어휘를 인식하는 고립단어 인식시스템의 경우에는 개인용 컴퓨터에서도 별도의 DSP 보드없이 실시간 처리가 가능하며, 둘째, 추출방법자체에서 음성의 크기 정보를 이용하지 않음으로써 별도의 에너지 정규화 과정이 필요 없고, 셋째, 파라미터 추출에 사용된 간격

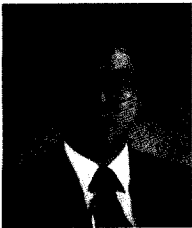
정보가 인지적 측면에서 개인성 정보가 유지되지 않기 때문에, 추출된 파라미터가 다른 파라미터에 비교하여 개인성에 따른 영향을 적게 받는다는 장점을 지닌다.

參 考 文 獻

- [1] 김순협, "국내외 음성 인식 기술 동향 및 전망," *Proc. of Korean-Japan Joint Workshop on Advanced Technology of Speech Recognition and Synthesis*, pp. 183-198, Jul. 1991.
- [2] T. Sakai and S. Inoue, "New Instruments and Methods for Speech Analysis," *J. Acoust. Soc. Am.* vol. 32, pp. 441-450, 1960.
- [3] T. Sakai and S. Doshita, "The Automatic Speech Recognition System for Conversational Sound," *IEEE Trans. on Electronic Computers*, vol. EC-12, pp. 835-846, Dec. 1963.
- [4] W. Bezdel and H.J. Chandler, "Results of an Analysis and Recognition of Vowels by Computer using Zero-Crossing Data," *Proc. IEE*, vol. 112, no. 11, pp. 2060-2066, Nov. 1965.
- [5] J.M. Baker, "A New Time-Domain Analysis of Human Speech and Other Complex Waveforms," Ph.D. Dissertation, Carnegie-Mellon Univ., Pittsburgh, PA., 1975.
- [6] J.C.R. Licklider and I.P. Pollack, "Effects of Differentiation, Integration, and Infinite Peak Clipping upon the Intelligibility of Speech," *J. Acoust. Soc. Am.* vol. 20, no. 1, pp. 42-51, Jan. 1948.
- [7] W.A. Ainsworth, "Relative Intelligibility of Different Transforms of Clipped Speech," *J. Acoust. Soc. Am.* vol. 41, pp. 1272-1276, 1967.
- [8] 임재열, 백성준, 안수길, 김범훈, "극대 및 극소점에서의 비균일 표본화에 의한 음성압축," *한국음향학회지*, 11권, 4호, pp. 36-44, 1992.
- [9] N.S. Jayant and P. Noll, *Digital*

- Coding of Waveforms: Principles and Applications to Speech and Video. Prentice-Hall, Inc., 1984.
- [10] L.R. Rabiner and M.R. Sambur. "An Algorithm for Determining the Endpoints of Isolated Utterances." *Bell Syst. Tech. J.*, vol. 54, no. 2, pp. 297-315, Feb. 1975.
- [11] B.A. Bladon and B. Lindblom. "Modelling the Judgement of Vowel Quality Difference." *J. Acoust. Soc. Am.* vol. 69, pp. 1414-1422, 1981.
- [12] Y. Linde, A. Buzo, and R.M. Gray. "An Algorithm for Vector Quantizer Design." *IEEE Trans. on Commun.*, vol. COM-28, pp. 84-95, Jan. 1980.
- [13] B.H. Juang, L.R. Rabiner, and J.G. Wilpon. "On the Use of Band Pass Liftering in Speech Recognition." *IEEE Trans. on ASSP*, vol. ASSP-35, pp. 947-953, 1989.
- [14] L.R. Rabiner, S.E. Levinson, and M. M. Sondhi. "On the use of Hidden Markov Models for Speaker-Independent Recognition of Isolated Words from a Medium Size Vocabulary." *AT&T Bell System Technical Journal*, vol. 63, pp. 627-641, 1984.
- [15] 이대룡. "VQ와 HMM을 이용한 한국어 음성 인식에서 화자적응에 관한 연구." 서울대학교 대학원 전자공학과 공학석사학위 논문, 1992.

 著者紹介



林材烈(正會員)

1964年 1月 15日生. 1986年 2月 서울대 전자공학과 졸업(공학사). 1988年 2月 서울대 대학원 전자공학과 졸업(공학석사). 현재 동대학원 전자공학과 박사과정. 주관심 분야는 통신이론, 음성신호처리.

적응신호처리 등임.



金亨一(學生會員)

1969年 11月 11日生. 1991年 2月 한국과학기술대학 전기 및 전자공학화 졸업(공학사). 1993年 2月 서울대 대학원 전자공학과 졸업(공학석사). 현재 동대학원 전자공학과 박사과정. 주관심 분야

는 추정이론 및 신호처리 등임.

安秀桔(正會員) 第24卷 第3號 參照

현재 서울대학교 전자공학과 교수